BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 6: e91168
doi: 10.3897/biss.6.91168

OPEN ACCESS

Conference Abstract

Zen and the Art of Persistent Identifier Service Development for Digital Specimen

Sharif Islam[‡,§], Soulaine Theocharides[‡,§], Wouter Addink[‡,§]

‡ Naturalis Biodiversity Center, Leiden, Netherlands
§ Distributed System of Scientific Collections - DiSSCo, Leiden, Netherlands

Corresponding author: Sharif Islam (sharif.islam@naturalis.nl)

Received: 02 Aug 2022 | Published: 02 Aug 2022

Citation: Islam S, Theocharides S, Addink W (2022) Zen and the Art of Persistent Identifier Service Development for Digital Specimen. Biodiversity Information Science and Standards 6: e91168.
https://doi.org/10.3897/biss.6.91168

Abstract

One of the most desired (and still missing) elements to enable the concept of Digital Extended (Webster et al. 2021) Specimens is a persistent identifier (PID) for the new digital specimen object. Digital Specimens are created to act as a digital surrogate of the physical objects. Digital Specimens contain all data relevant to the specimens as well as derived data like genetic sequences, trait information, and references to publications, species and environmental information. A PID for the Digital Specimen is thus essential to link it to the extended information. Furthermore, the extended information also needs PIDs to create unique and resolvable references to enable bidirectional linking.

DiSSCo, Distributed System of Scientific Collections, has done extensive work to select the most appropriate PID scheme (Hardisty et al. 2021) and design a PID infrastructure for the pan-European specimen collections. The design was created in the Biodiversity Community Integrated Knowledge Library project (BiCIKL) with funding from the European Commission. The draft design has been discussed with technical specialists in the joint DiSSCo and Consortium of European Taxonomic Facilities (CETAF) community, with international stakeholders like GBIF and iDigBio. The draft was also presented at the 2022 conference of SPNHC (the Society for the Preservation of Natural History Collections) for further feedback, after which it was finalised.

© Islam S et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

This talk will provide a short overview of the key elements in the design of the Persistent Identifier system for Digital Specimens such as the metadata schema, a human friendly string format, and PID lifecycle with support for state changes of the physical object such as splitting and merging. We will illustrate this by showing the current status of development.

The Digital Specimen forms the backbone of the FAIR (Findable, Accessible, Interoperable, Reusable) services (Addink et al. 2021) in development by the DiSSCo Prepare project, a project to prepare for the new research infrastructure which currently involves 170 institutions in 23 countries holding an estimated 1,5 billion specimens. The Digital Specimen is a FAIR Digital Object with a PID and machine actionable metadata to make it FAIR. The PID service ensures the identifiers' integrity and preserves these linkages, improving the FAIR quality of all specimen data. The talk will describe the system design and current test deployment setup, as well as delve into a few technical details of the Handle Server architecture, PID Kernel records, and a roadmap for the service development. The design and test work are done under the auspices of the BiCIKL project that aims to foster collaboration between infrastructures and develop bidirectional connections between the data so that they become part of a big FAIR data pool seamlessly available to researchers (Penev et al. 2022).

Keywords

FAIR, FAIR digital objects, DiSSCo, BiCIKL, natural science collections, PID, infrastructure development

Presenting author

Soulaine Theocharides

Presented at

TDWG 2022

Funding program

Horizon 2020

Grant title

DiSSCo Prepare (Grant agreement no. 871043)

BiCIKL (Grant agreement no. 101007492)

References

- Addink W, Islam S, Alonso J (2021) DiSSCo e-Services to Serve Global Community Needs. Biodiversity Information Science and Standards 5 https://doi.org/10.3897/biss.5.73903
- Hardisty A, Addink W, Glöckler F, Güntsch A, Islam S, Weiland C (2021) A choice of persistent identifier schemes for the Distributed System of Scientific Collections (DiSSCo). Research Ideas and Outcomes 7 https://doi.org/10.3897/rio.7.e67379
- Penev L, Koureas D, Groom Q, Lanfear J, Agosti D, Casino A, Miller J, Arvanitidis C, Cochrane G, Hobern D, Banki O, Addink W, Kõljalg U, Copas K, Mergen P, Güntsch A, Benichou L, Benito Gonzalez Lopez J, Ruch P, Martin C, Barov B, Demirova I, Hristova K (2022) Biodiversity Community Integrated Knowledge Library (BiCIKL). Research Ideas and Outcomes 8 https://doi.org/10.3897/rio.8.e81136
- Webster MS, Buschbom J, Hardisty A, Bentley A (2021) The Digital Extended Specimen will Enable New Science and Applications. *Biodiversity Information Science and Standards* 1 https://doi.org/10.3897/biss.5.75736